

UC Riverside

UC Riverside Previously Published Works

Title

Draft Genome Sequence of the Yeast *Rhodotorula* sp. Strain CCREE 5036, Isolated from McMurdo Dry Valleys, Antarctica.

Permalink

<https://escholarship.org/uc/item/7395q9j5>

Journal

Microbiology resource announcements, 9(14)

ISSN

2576-098X

Authors

Coleine, Claudia
Masonjones, Sawyer
Onofri, Silvano
[et al.](#)

Publication Date

2020-04-01

DOI

10.1128/mra.00020-20

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



Draft Genome Sequence of the Yeast *Rhodotorula* sp. Strain CCREE 5036, Isolated from McMurdo Dry Valleys, Antarctica

 Claudia Coleine,^a Sawyer Masonjones,^b Silvano Onofri,^a Laura Selbmann,^{a,c}  Jason E. Stajich^b

^aDepartment of Ecological and Biological Sciences, University of Tuscia, Viterbo, Italy

^bDepartment of Microbiology and Plant Pathology, University of California—Riverside, Riverside, California, USA

^cItalian National Antarctic Museum (MNA), Mycological Section, Genoa, Italy

ABSTRACT A draft genome sequence was assembled and annotated of the basidiomycetous yeast *Rhodotorula* sp. strain CCREE 5036, isolated from Antarctic soil communities. The genome assembly is 19.07 megabases and encodes 6,434 protein-coding genes. The sequence will contribute to understanding the diversity of fungi inhabiting polar regions.

Rhodotorula fungi are ubiquitous saprophytic yeasts taxonomically classified in the Pucciniomycotina and Ustilaginomycotina subphyla (phylum Basidiomycota) (1, 2). These fungi can be isolated from many environments and are often found associated with humans, animals, and food (3). Species have been described from the gut microbiota of carnivorous fish (4) and contaminated soil (5). Some members of this group are cryophilic extremophiles and can persist under extreme conditions (low temperature, high salinity, high pressure, and low pH) (6–11). The genome sequence of an Antarctic *Rhodotorula* isolate will be useful for comparative studies of evolution of extremophilic yeasts, in efforts to study their role in biogeochemical nutrient cycling in cold environments, and in bioprospecting for new enzymes (12, 13).

A *Rhodotorula* sp. culture was isolated from soil collected near a glacier during the XI Italian Antarctic Expedition (1995 to 1996) at Edmonson Point at 74°20'00"S, 165°08'00"E (Northern Victoria Land, Continental Antarctica), an Antarctic Specially Protected Area (ASPA), following the protocol described by Selbmann et al. (14). Briefly, soil was sprinkled on petri dishes containing 2% malt extract agar (MEA; AppliChem GmbH, Darmstadt, Germany) supplemented with 100 ppm chloramphenicol and incubated at 10°C for several months. Yeast colonies were streaked onto fresh medium to isolate pure cultures. *Rhodotorula* sp. CCREE 5036 strain culture is deposited in the Culture Collection of Fungi from Extreme Environments (CCREE; University of Tuscia, Italy) and at the Dipartimento di Biologia Vegetale e Agroambientale of the University of Perugia Industrial Yeasts Collection (DBVPG) as strain 5527. Genomic DNA was extracted from a pure culture grown for 3 weeks at 10°C on MEA following the cetyltrimethylammonium bromide (CTAB) protocol (15). The DNA was sheared with a Covaris S220 ultrasonicator, and a sequencing library was constructed using the Neoprep TruSeq nano DNA sample prep protocol (Illumina, Inc., San Diego, CA) in a genomics core (Institute for Integrative Genome Biology, University of California, Riverside). The library was multiplexed and sequenced on an Illumina MiSeq flow cell to obtain 6.1 million 2 × 300-bp paired-end sequence reads. FastQC (v0.11.3) was used to check read quality (16).

Genome assembly was performed with MaSuRCA (v2.3.2) (17) using default parameters (cgwErrorRate, 0.15), which included quality-based read trimming and corrections. Trimmed reads averaged 199 bp. Assembled scaffolds were filtered for vector contamination with Sequin (v15.10) (<https://www.ncbi.nlm.nih.gov/Sequin/>), and redundant

Citation Coleine C, Masonjones S, Onofri S, Selbmann L, Stajich JE. 2020. Draft genome sequence of the yeast *Rhodotorula* sp. strain CCREE 5036, isolated from McMurdo Dry Valleys, Antarctica. Microbiol Resour Announc 9:e00020-20. <https://doi.org/10.1128/MRA.00020-20>.

Editor Antonis Rokas, Vanderbilt University

Copyright © 2020 Coleine et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Laura Selbmann, salbmann@unitus.it, or Jason E. Stajich, jason.stajich@ucr.edu.

Received 9 January 2020

Accepted 14 March 2020

Published 2 April 2020

scaffolds were eliminated if they aligned with at least 95% identity to a longer contig with MUMmer (v3.23) (18), using the “clean” step in Funannotate (v0.5.5) (19). The assembly was 155 contigs and totaled 19.08 Mb in length (N_{50} , 338 kb; L_{50} , 19; longest scaffold, 930,366 bp; G+C content, 60.58%; average depth of coverage, $192\times$).

Genome annotation performed by Funannotate (v0.5.5) (19) produced consensus gene models by EVIDENCEModeler (EVM) (20), combining *ab initio* predictions from AUGUSTUS (v3.2.2) (21) and GeneMark.hmm-ES (v4.32) (22) with protein-to-genome alignments from Exonerate (v2.2.0) (23). GeneMark.hmm-ES self-training used default parameters, and AUGUSTUS was trained with alignments of BUSCO basidiomycota_odb9 proteins (v9) (24) and gene prediction parameters archived in a GitHub repository (25). Gene functions were assigned by similarity to Pfam (26), MEROPS (27), CAZy (28, 29), eggNOG (v4.5) (30), InterProScan (31), and Swissprot (32) databases by BLASTP (v2.5.0+) or HMMER3 (33) searches using Funannotate default parameters. A total of 6,553 protein-coding genes were predicted and prepared for GenBank submission by Genome Annotation Generator (34).

Data availability. This whole-genome shotgun project was deposited at DDBJ/ENA/GenBank under the accession number [MXAQ00000000](#). The version described in this paper is the first version, MXAQ01000000. Illumina sequence reads are released under SRA accession number [SRR5223778](#) and associated with BioProject [PRJNA342238](#).

ACKNOWLEDGMENTS

The Italian Antarctic National Museum (MNA) is kindly acknowledged for financial support to the Mycological Section on the MNA and for providing the strains sequenced in this study that are stored in the Culture Collection of Fungi from Extreme Environments (CCFEE) (University of Tuscia, Italy). J.E.S. is a CIFAR fellow in the Fungal Kingdom: Threats and Opportunities program. C.C. and L.S. kindly acknowledge the Italian National Program for Antarctic Researches (PNRA) for funding sampling campaigns and the research activities in Italy. Sequencing was supported through United States Department of Agriculture, National Institute of Food and Agriculture Hatch project CA-R-PPA-5062-H to J.E.S. Data analyses were performed on the High-Performance Computing Cluster at the University of California, Riverside, in the Institute of Integrative Genome Biology, supported by NSF DBI-1429826 and NIH S10-OD016290.

We declare no competing interests.

REFERENCES

- Aime MC, Toome M, McLaughlin DJ. 2014. Pucciniomycotina, p 271–294. In *Systematics and evolution*. Springer, Berlin, Germany.
- Urbina H, Aime MC. 2018. A closer look at Sporidiobolales: ubiquitous microbial community members of plant and food biospheres. *Mycologia* 110:79–92. <https://doi.org/10.1080/00275514.2018.1438020>.
- Larone DH. 2011. *Medically important fungi: a guide to identification*, 5th ed. ASM Press, Washington, DC.
- Raggi P, Lopez P, Diaz A, Carrasco D, Silva A, Velez A, Opazo R, Magne F, Navarrete PA. 2014. *Debaryomyces hansenii* and *Rhodotorula mucilaginosa* comprised the yeast core gut microbiota of wild and reared carnivorous salmonids, croaker and yellowtail. *Environ Microbiol* 16: 2791–2803. <https://doi.org/10.1111/1462-2920.12397>.
- Chandran P, Das N. 2012. Role of plasmid in diesel oil degradation by yeast species isolated from petroleum hydrocarbon-contaminated soil. *Environ Technol* 33:645–652. <https://doi.org/10.1080/09593330.2011.587024>.
- Feller G, Gerday C. 2003. Psychrophilic enzymes: hot topics in cold adaptation. *Nat Rev Microbiol* 1:200–208. <https://doi.org/10.1038/nrmicro773>.
- Margesi R, Fonteyne PA, Schinner F, Sampaio JP. 2007. *Rhodotorula psychrophila* sp. nov., *Rhodotorula psychrophilica* sp. nov. and *Rhodotorula glacialis* sp. nov., novel psychrophilic basidiomycetous yeast species isolated from alpine environments. *Int J Syst Evol Microbiol* 57:2179–2184. <https://doi.org/10.1099/ijs.0.65111-0>.
- Bergauer P, Fonteyne PA, Nold N, Schinner F, Margesi R. 2005. Biodegradation of phenol and phenol-related compounds by psychrophilic and cold-tolerant alpine yeasts. *Chemosphere* 59:909–918. <https://doi.org/10.1016/j.chemosphere.2004.11.011>.
- Goordial J, Raymond-Bouchard I, Riley R, Ronholm J, Shapiro N, Woyke T, LaButti KM, Tice H, Amirebrahimi M, Grigoriev IV, Greer C, Bakermans C, Whyte L. 2016. Improved high-quality draft genome sequence of the eurypsychrophile *Rhodotorula* sp. JG1b, isolated from permafrost in the hyperarid upper-elevation McMurdo Dry Valleys, Antarctica. *Genome Announc* 4:e00069-16. <https://doi.org/10.1128/genomeA.00069-16>.
- Burgaud G, Arzur D, Durand L, Cambon-Bonavita M-A, Barbier G. 2010. Marine culturable yeasts in deep-sea hydrothermal vents: species richness and association with fauna. *FEMS Microbiol Ecol* 73:121–133. <https://doi.org/10.1111/j.1574-6941.2010.00881.x>.
- Sannino C, Tasselli G, Filippucci S, Turchetti B, Buzzini P. 2017. Yeasts in nonpolar cold habitats, p 367–396. In *Yeasts in natural ecosystems: diversity*. Springer, Cham, Switzerland.
- Raspor P, Zupan J. 2006. Yeasts in extreme environments. In Peter G, Rosa C, ed, *The Yeast handbook*. Springer, Cham, Switzerland.
- Foght J, Aislabie J, Turner S, Brown CE, Ryburn J, Saul DJ, Lawson W. 2004. Culturable bacteria in subglacial sediments and ice from two Southern Hemisphere glaciers. *Microb Ecol* 47:329–340. <https://doi.org/10.1007/s00248-003-1036-5>.
- Selbmann L, Zucconi L, Onofri S, Cecchini C, Isola D, Turchetti B, Buzzini P. 2014. Taxonomic and phenotypic characterization of yeasts isolated from worldwide cold rock-associated habitats. *Fungal Biol* 118:61–71. <https://doi.org/10.1016/j.funbio.2013.11.002>.
- Fulton TM, Chunwongse J, Tanksley SD. 1995. Microprep protocol for

- extraction of DNA from tomato and other herbaceous plants. *Plant Mol Biol Rep* 13:207–209. <https://doi.org/10.1007/BF02670897>.
16. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
 17. Zimin AV, Marçais G, Pui D, Roberts M, Salzberg SL, Yorke JA. 2013. The MaSuRCA genome assembler. *Bioinformatics* 29:2669–2677. <https://doi.org/10.1093/bioinformatics/btt476>.
 18. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. *Genome Biol* 5:R12. <https://doi.org/10.1186/gb-2004-5-2-r12>.
 19. Palmer J, Stajich JE. 2017. Funannotate: eukaryotic genome annotation pipeline. <https://doi.org/10.5281/zenodo.1134477>.
 20. Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. 2008. Automated eukaryotic gene structure annotation using EVIDENCEModeler and the Program to Assemble Spliced Alignments. *Genome Biol* 9:R7. <https://doi.org/10.1186/gb-2008-9-1-r7>.
 21. Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. 2006. AUGUSTUS: *ab initio* prediction of alternative transcripts. *Nucleic Acids Res* 34:W435–W439. <https://doi.org/10.1093/nar/gkl200>.
 22. Lomsadze A, Ter-Hovhannisyan V, Chernoff YO, Borodovsky M. 2005. Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res* 33:6494–6506. <https://doi.org/10.1093/nar/gki937>.
 23. Slater GSC, Birney E. 2005. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6:31. <https://doi.org/10.1186/1471-2105-6-31>.
 24. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>.
 25. Stajich JE. 2018. hyphalip/fungi-gene-prediction-params: fungi gene prediction set v.0.1.0. <https://doi.org/10.5281/zenodo.1649679>.
 26. Finn RD, Bateman A, Clements J, Coghill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M. 2014. Pfam: the protein families database. *Nucleic Acids Res* 42:D222–D230. <https://doi.org/10.1093/nar/gkt1223>.
 27. Rawlings ND, Barrett AJ, Finn R. 2016. Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* 44:D343–D350. <https://doi.org/10.1093/nar/gkv1118>.
 28. Huang L, Zhang H, Wu P, Entwistle S, Li X, Yohe T, Yi H, Yang Z, Yin Y. 2018. dbCAN-seq: a database of carbohydrate-active enzyme (CAZyme) sequence and annotation. *Nucleic Acids Res* 46:D516–D521. <https://doi.org/10.1093/nar/gkx894>.
 29. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42:D490–D495. <https://doi.org/10.1093/nar/gkt1178>.
 30. Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M, Jensen LJ, von Mering C, Bork P. 2016. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* 44:D286–D293. <https://doi.org/10.1093/nar/gkv1248>.
 31. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30:1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>.
 32. Boutet E, Lieberherr D, Tognolli M, Schneider M, Bansal P, Bridge AJ, Poux S, Bougueleret L, Xenarios I. 2016. UniProtKB/Swiss-Prot, the manually annotated section of the UniProt KnowledgeBase: how to use the entry view. *Methods Mol Biol* 1374:23–54. https://doi.org/10.1007/978-1-4939-3167-5_2.
 33. Eddy SR. 2011. Accelerated profile HMM searches. *PLoS Comput Biol* 7:e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>.
 34. Hall B, DeRego T, Geib S. 2014. Genome Annotation Generator: a simple tool for generating and correcting WGS annotation tables for NCBI submission. *GigaScience* 7:1–5. <https://doi.org/10.1093/gigascience/giy018>.